ORIGINAL RESEARCH PAPER

Engineering

KEY WORDS: Intelligent word recognition, Character conversion, Digit recognition, Machine Learning

CONVERSION OF HANDWRITTEN DATA INTO DIGITAL VERSION

Neha Tiwari — Assistant Professor, Department of Electronics and Telecommunication, Dr. D. Y. Patil Institute of Engineering Management and Research Akurdi, Pune, Maharashtra, India.

Sushant Narawade* — Student, Department of Electronics and Telecommunication, Dr. D. Y. Patil Institute of Engineering Management and Research Akurdi, Pune, Maharashtra, India.*Corresponding Author

Pratik Chavan — Student, Department of Electronics and Telecommunication, Dr. D. Y. Patil Institute of Engineering Management and Research Akurdi, Pune, Maharashtra, India.

Omkar Kolhe — Student, Department of Electronics and Telecommunication, Dr. D. Y. Patil Institute of Engineering Management and Research Akurdi, Pune, Maharashtra, India.

ABSTRACT

In handwritten recognition, we are facing the issue concerning to the recognition of words. Because there are various types of handwriting style. This paper presents an approach to recognition of handwritten words and digit based on machine learning technique. The ability to develop an efficient algorithm that can recognize handwritten words and digits. Which is submitted by users by the way of a scanner and other digital devices. The paper presents a new application that can turn handwritten page into computerized version by using Intelligent word recognition (IWR). It will be useful friendly to read the ancient script.

I. INTRODUCTION

This paper refers to the system which convert the handwritten text or a conversation into a machine coded language, text document or scanned images etc. As we know the handwritten character recognition is a bit difficult job to do because of different variation of the writing styles. Identification of various style of text using different kind of method is one of the key features in the area of text recognition. Because of the wide range of writing style. The system must be robust to improve the extraction and performance of the system. Nowadays the handwritten text recognition has gained a lot of interests of the industries sowing its application in various fields.

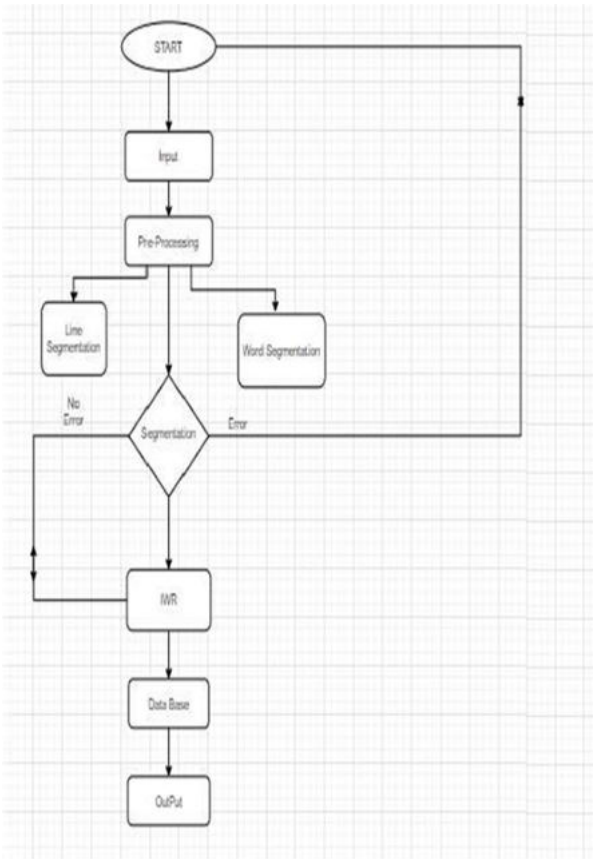
This system uses the Intelligent Word Recognition (IWR) technology which makes the system less complex and makes the process easier to recognize the meaningful word or sentences. Because of these technologies the digitization of the text has become simpler. The group of neurons helps capturing the pixel of the input image of the text and extracts a meaning full text or digit by using the intelligent word recognition. Further these data are passed on and transformed into a meaningful manner according to the designed network. Again, the different processes are done to extraction/receive the expected output or show that character was read at the input side.

II. LITERATURE REVIEW

Sr. No.	Title of the paper	Year of publication	Authors	Methodology
1	Handwritten Document into Digitized Text Using Segmentation Algorithm	2014	D. Kavitha, P. Shamini	1. Data collection done using Kaggle dataset 2. Handwritten text is digitized using segmentation algorithm 3. Intelligent Word Recognition (IWR) technique is used which uses neural network to find meaningful words 4. Segmentation experiments were done by ANN trained segmentation algorithm 5. Sample of handwriting texts from database used
2	Handwritten Optical Character Recognition (OCR): A Comprehensiv e Systematic Literature Review (SLR)	2020	Jamshed Memon, Maira Sami, Rizwan Ahmed Khan	1. Optical Character Recognition (OCR) is used 2. English, Urdu, Arabic, Indian, Farsi, Chinese language is tested 3. Different algorithm is used to different languages 4. Good accuracy obtained on Devanagari & Bangla numerals by multilayer perceptron
3	Handwritten Digit Recognition using Machine Learning Algorithms	2018	S.M.Shamim, Angona Sarker, Md Badrul Alam Miah	1. Only digit recognition is done 2. Machine learning algorithms is used 3. Higher accuracy achieved which is 90.37%
4	Handwriting Recognition of Historical Documents with few labeled data	2018	Edgard Chammas and Chafic Mokbel	1. Convolutional recurrent neural network (CRNN) trained deep data only 10% manually written text line 2. Obtained second rank result in ICDAR2017 competition 3. Results can be more accurate by improving efficiency of segmentation algorithm

5	Effective handwriting Digit Recognition using deep convolution neural network	2020	Yellapragada SS Bharadwaj, Rajaram P, Sriram V.P, Sudhakar S	1. Digit recognition is done. 2. A simple network approach towards handwritten digit recognition using convolution. 3. Accuracy 98.51% in this method
6	Effective Handwritten Digit Recognition Based on Multi-feature Extraction and deep analysis	2015	Caiyun Ma & Hong Zhang	1. Digit Recognition in computer vision 2. Based on multi-feature extraction and deep analysis 3. MNIST Data set is used

FLOW CHART:



III. METHODOLOGY

1- Data Collection

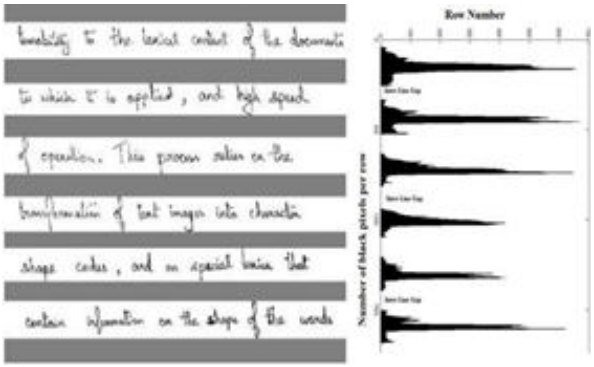
- Using data base collected from Kaggle, UCI website.

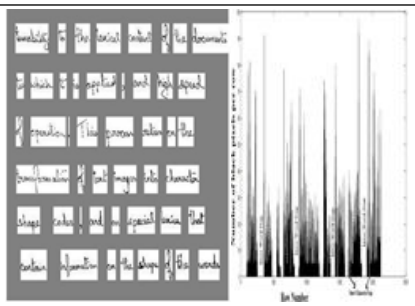
2- Input

- Upload an image that we have to convert.
- Check there is any error in the program. If there is no error then upload the input image.

3 Pre-Processing

- Line segmentation check every line by line.
- Word segmentation check every word and symbol.





4- Segmentation

- In this step were used to extracting the character through words.
- Segmentation algorithm share out with the handwritten image and take-out separate characters from handwritten words.
- This transforms into digitized text through machine learning.

5 Recognition engine

- IWR engine understands the transform words and allow whether it has meaning or not.
- In this can achieve online conversion through proposed system.

6- Data Base

- Check the words from data base in digital version.

7- Post Processing

- Digitized words are from after recognizing Analyze words with semantic meanings.

IV. CONCLUSION

The main perspective of this project is to recognition of handwritten text-lines with the better accuracy. The main focus of this investigation is to improve the system to identify any other character and tried to make our system more font independent that our system can read or recognize any other font. An Intelligent Word Recognition (IWR) and the segmentation algorithm techniques has obtained good results.

V. FUTURE SCOPE

In this project, this work can be expanding in future to recognition of different languages Characters. So, it will be able to convert the historical texts or documents which will much helpful to us to know ancient history. It will also helpful to read old dull government documents. It can be used to read address written on post in post office.

REFERENCES

[1] D. Kavitha, P. Shamini, "Handwritten Document into Digitized Text Into Using Segmentation Algorithm". Department of Computer Application, Easwari Engineering College, Tamil nadu (ISSN-2014)
[2] Jamshed Memon, Maira Sami, Rizwan Ahmed Khan. "Handwritten Optical Recognition (OCR): A Comprehensive Systematic Literature Review (SLR)". Faculty of IT, Barrett Hodgson University 1 January, 2020
[3] S M Shamim, Mohammad Badrul Alam Miah, Angona Sarker. "Handwritten Digit Recognition Using Machine Learning Algorithms". Global Journal of Computer Science and Technology.
[4] Edgard Chammas & Chafic Mokble. "Handwriting Recognition of Historical Documents with few Lebeled Data". University of balamand El-koura, Lebanon
[5] Yellapragada SS Bharadwaj, Rajaram P, Sriram V.P, "Effective Handwritten

Digit Recognition using Deep Convolution Neural Network." Dept. of CSE koneru lakshmaish education foundation, India.
[6] Caiyun Ma & Hong Zhang, "Effective Handwritten Digit Recognition Based on Multi- feature Extraction and Deep Analysis." College of computer science and technology, Wuhan
[7] Bhowmik S, Roushan M. G, Sarkar R, Nasipuri M, Polley S. & Malakar S. "Handwritten Bangla Word Recognition Using HOG Descriptor." In 2014 Fourth International Conference of Emerging Applications of Information Technology (EAIT). IEEE, 2014
[8] Kruse R, Borgelt C, Klawonn F, Moewes C, Steinbrecher M & Held P. Multi-layer Perceptrons. In Computational Intelligence (pp. 47-81). Springer London 2013.